FIGURE 1 diagrammatically illustrates a speech engine in accordance with this invention; and

FIGURE 2 shows a speech engine as illustrated in Figure 1 attached to a telephone network.

DETAILED DESCRIPTION OF EXEMPLARY EMBODIMENTS--.

Page 9, line 15, after "P3" insert a comma.

Page 12, line 2-8, delete entirely.

IN THE CLAIMS

To address formality objections, please amend the format of claims 1-4:

1. (Amended) A method of converting an input signal [into an output signal, wherein said input signal represents] representing a text in phonemes [and said] into an output [signal is a] digital waveform signal convertible into an acoustic waveform corresponding to said text, wherein said method comprises:[-]

(a) dividing said input signal into input segments, each of which is stored in [the] an output section of a linked database [,];

(b) for each input segment identified in step (a), retrieving an output [a] segment of said digital waveform from [the] an output section of the database, said output segment being that which is linked to the input segment[,] ; and

(c) joining the digital output segments retrieved in step (b), said output segments being kept in the same order as the [equivalent] respectively associated input segments[,] whereby the resulting output digital signal is a waveform corresponding to the input signal [, characterised in that] waveform;

the output section of the database [contains] containing an extended digital waveform containing plural contextual occurrences in extended speech representing signals of the phonemes to be converted and having a location parameter for identifying any point therein whereby the establishment of beginning and ending location parameters defines a portion of said extended digital waveform [, and] ;

step (a) [comprises] includes establishing beginning and ending location parameters for segments of the input signal; and

step (c) [comprises utilising] includes utilizing the parameters established in step (a) for retrieving a portion of stored digital waveform.

2. (Amended) A method [according to] as in claim 1, wherein step (a) comprises:

comparing windows of input signal with windows of the [input] access section of the database to establish a closest match for the input signal.

3. (Amended) A method [according to] as in claim 2, wherein each window of input signal has a length equivalent to [5] five phonemes.

4. (Amended) A method [according to] as in claim 3, in which:

the [input] access section of the database is [organised] organized into three hierarchical levels [; namely] comprising:

(i) a top level containing single phonemes corresponding to the [central] third phoneme of a window;

(ii) a second level which contains [the] equivalents of the second and fourth phonemes of a window;

(iii) a lowest level which contains [the] equivalents of the first and fifth phonemes of the window, whereby identification of a portion of the lowest level identifies a stored window of phonemes; and

[and] the [matching] comparing step comprises:

selecting an exact match for the [central] third phoneme of the input window from the [first] top level of the hierarchy,

selecting a best match for the second and fourth phonemes [2 and 4] from the second level of the hierarchy corresponding to the earlier selected portion of the top level of the hierarchy and,

finally, selecting from the bottom level of the hierarchy [the] a best match for the first and fifth phonemes [1 and 5] from that portion of the lowest [bottom] level which corresponds to the earlier selection in the second level of the hierarchy.

Add the following new claims 11 through 17:

--11. A method of converting an input signal into an output signal, wherein:

(a) said input signal represents a text in phonemes;

(b) said output signal is a digital waveform convertible into an acoustic waveform corresponding to said text;

(c) a database is used having an input section and an output section;

(d) said output section containing an extended digital waveform having a location parameter for identifying any point therein whereby the

establishment of beginning and ending location parameters defines a portion of said extended digital waveform;

(e) said input section containing segments of an extended phoneme text corresponding to the extended waveform contained in the output section;

said method comprising the steps of:

(i) dividing said input signal into input segments;

(ii) matching said input segments with segments contained in the input section of the database thereby establishing beginning and ending location parameters;

(iii) retrieving from the output section of said database segments of extended digital waveform corresponding to said beginning and ending location parameters; and

(iv) joining the output segments of digital waveform so retrieved, said segments being kept in the same order as the corresponding input segments.

12. A method of converting an input signal into an output signal, wherein:

(a) said input signal represents an input text in phonemes;

(b) said output signal is a digital waveform convertible into an acoustic waveform corresponding to said input text;

(c) a database is used having an input section and an output section;

(d) said output section containing an extended digital waveform having a location parameter for identifying any point therein whereby the establishment of beginning and ending location parameters defines a portion of said extended digital waveform;

(e) said input section defining context windows of an extended phoneme text corresponding to the extended waveform contained in the output section;

said method comprising the steps of:

(i) dividing said input signal into input segments;

(ii) matching said input segments with context windows contained in the input section of the database thereby establishing beginning and ending location parameters;

(iii) retrieving from the output section of said database segments of extended waveform corresponding to said beginning and ending location parameters; and

(iv) joining the output segments of a digital waveform, said joined segments being kept in the same order as the corresponding input segments.

13. A method as in claim 12 wherein each context window has a length equivalent to five phonemes.

14. A method as in claim 13 in which:

the context windows are stored in three hierarchical levels comprising:

(i) a top level containing single phonemes corresponding to the third phoneme of a window;

(ii) a second level which contains equivalents of the second and fourth phonemes of a window; and

(iii) a lowest level which contains equivalents of the first and fifth phonemes of the window, whereby identification of a portion of the lowest level identifies a stored window of phonemes; and

the matching step comprises:

selecting an exact match for the third phoneme of the input window from a first level of the hierarchy,

selecting a best match for the second and fourth phonemes from a second level of the hierarchy corresponding to the earlier selected portion of the top level of the hierarchy and,

finally, selecting from the lowest level of the hierarchy a best match for the first and fifth phonemes from that portion of the lowest level which corresponds to the earlier selection in the second level of the hierarchy.

15. A method of converting a string of input phoneme text signals into an output digital waveform signal representing acoustic speech, said method comprising the steps of:

(a) storing extended digital speech waveform signals, representing plural utterances of each phoneme to be converted, in a corresponding plurality of speech contexts with different preceding and/or succeeding phonemes;

(b) dividing an input string of phonemes into input subsets of N contiguous phonemes, N being an integer;

(c) matching each said input subset with a most similar corresponding subset of N contiguous phonemes in said stored extended digital speech waveform;

(d) selecting a portion of the stored extended digital speech waveform corresponding to at least one phoneme of the matched subset; and

repeating at least steps (c) and (d) while concatenating the thus-selected portions of the extended digital speech waveform to provide said converted output digital waveform signal representing acoustic speech.

16. A method as in claim 16 wherein N equals five.

17. A method as in claim 16 wherein:

N equals an odd integer equal to three or greater and wherein a hierarchical database is maintained with:

(i) a top level containing single phonemes corresponding to the center or (N+1)/2 phoneme of each subset;

(ii) at least one lower level containing plural phonemes that are contiguous to the center phoneme of each subset; and